

Fileservice mit ADSM

Da wir mit unserem VHS-Roboter und der UniTree-Kapazität am oberen Ende angekommen waren, wurden für uns größere Umstellungen bzw. Neuanschaffungen erforderlich. Aus diesem Grund wurde auch das Konzept des Fileservices an der Humboldt-Universität neu überdacht.

In erster Linie ging es um das Verhältnis zwischen der Nutzung des HSM-Systems (*Hierarchical Storage Management*) UniTree und dem Backupsystem Sm-arch. Eine Analyse zeigte, dass auch UniTree von den Nutzern primär für die Sicherung ihrer Daten genutzt wurde.

Aus diesem Grund richteten wir unser Hauptaugenmerk auf ein System, das schwerpunktmäßig die Datensicherung ermöglicht. Weiterhin sollte es Möglichkeiten für eine längerfristige Aufbewahrung von Daten (Archivierung) und auch eine HSM-Komponente besitzen. Eine weitere Anforderung an das System war, dass die Datensicherung nicht nur native Unix Filesysteme unterstützt, sondern auch ein Backup von DCE/DFS ermöglicht, in dessen Einführungsphase wir gerade stehen. Durch das letzte Kriterium blieb für uns nach Sichtung mehrerer möglicher Systeme nur noch **ADSM** übrig.

Hier ein kleiner Einschub zur Namensgebung des Systems: **ADSM** (*Adstar Distributed Storage Manager*) ist ein IBM-Produkt, das während der Zeit der Beschaffung von Tivoli – einer IBM-Tochter – übernommen wurde. Dabei wurde ADSM zu TSM (Tivoli Storage Manager). Es werden also gewiss in den nächsten Zeiten die Begriffe ADSM und TSM gleichbedeutend benutzt.

Was liefert TSM nun für unsere Nutzer:

- **Backup** für eine Vielzahl von Rechnern (Unix, NT, MAC) der verschiedensten Hersteller und mit den verschiedensten Betriebssystemversionen (gewiss wird man bei heftigem Suchen eine Plattform finden, die nicht unterstützt wird)
- Aufheben mehrerer Versionen eines gesicherten Files – bei uns ist der Standard vier Versionen, von denen die alten Versionen nach vier Wochen entfernt werden. Die letzte Version des Files bleibt nach dem Löschen auf dem Client noch acht Wochen erhalten. Für ein File, der jeden Tag modifiziert wird, bedeutet das, dass die älteste Version 4 Tage alt ist.
- **Restore** ist in die Hand des Clients gegeben. Es bietet einen ausgefeilten Schutz der gesicherten Daten vor unberechtigtem Zugriff – grundsätzlich hat nur der Eigentümer der Daten Zugriff und das auch nur von dem Rechner aus, von dem gesichert wurde. Es gibt die Möglichkeit, dass man einem Nutzer (z. B.

root) auf einem anderen Rechner den Zugriff gestatten kann. Diese Möglichkeit sollte angewendet werden, da dadurch gewährleistet ist, dass man das Backup seiner Daten erreichen kann, auch wenn der eigene Rechner nicht funktionsfähig ist.

- Anstoßen der Sicherung vom Server aus (über ein dort aufgesetztes Scheduling) oder vom Nutzer (z. B. crontab oder direkt mittels Zeilenkommandos oder GUI); bei uns wird grundsätzlich das Scheduling vom Server aus eingesetzt (Ausnahmen bestätigen die Regel).
- **HSM**, vergleichbar mit der bisherigen UniTree-Funktionalität, an zwei Standorten – Physik und Rechenzentrum.

Hinweis zur Nutzung der HSM-Komponente:

Bei der Nutzung der HSM-Komponente sollten bedacht werden, dass das zugrunde liegende Filesystem ein Unix-System ist. Dies kann zur Folge haben, dass auf Daten, die von anderen Architekturen kommen, eventuell nicht mehr auf gewohnte Weise zugegriffen werden kann (Sonderzeichen in Filenamen!).

- **Archivierung** von Nutzerdaten, die sich zur Zeit nicht ändern. Es können Daten längerfristig aufgehoben werden. Die Daten können auf dem Quellrechner gestrichen werden, so dass der Nutzer dort wieder Platz gewinnt. Geplant sind zwei mögliche Archivierungszeiträume, die zwei bzw. fünf Jahre betragen. Die genaueren Modalitäten zur Archivierung werden später veröffentlicht.

Wie arbeitet TSM bei uns:

- Es wurden zwei TSM-Server installiert, die in den Räumen des Rechenzentrums in Mitte bzw. Adlershof stehen.
 - Mitte:
 - IBM H70 mit knapp 300 GB Plattenkapazität;
 - IBM Library 3494 (ca. 20 TB Kapazität)
 - Adlershof:
 - IBM H70 mit knapp 300 GB Plattenkapazität;
 - StorageTek Library 9710 (ca. 20 TB)
- Beide Server sind miteinander verbunden und bedienen jeweils die Clients, die sich in ihrer Umgebung befinden.
- Die Datensicherung findet nachts zwischen 0 Uhr und ca. 6 Uhr statt. Dabei werden die Daten in sogenannte Storagepools auf Platte gespeichert. Files, die größer als 100 MB sind, kommen sofort in die Library auf Tape.
- Am Tag werden alle Daten von Platte auf Tape migriert und von Tape aus zum jeweils anderen Standort (Mitte/Adlershof) kopiert.

Diese Vorgehensweise bietet in punkto Datensicherheit zwei Vorteile :

- man weiss sofort, dass die Daten auf den Tapes lesbar sind
- man hat eine zweite Kopie der Daten an einem entfernten Standort

Wie kann man TSM nutzen:

- In die Datensicherung werden von uns zurzeit Unix-Rechner und NT-Server einbezogen. Arbeitsgruppen sollten ihre zu sichernden Daten möglichst auf einigen wenigen Rechnern vorhalten, die dann von uns gesichert werden können. Dies ist notwendig, da die einzelnen Clients lizenziert werden müssen.

Eine Anforderung kann man durch eine Mail an *adsm@rz.hu-berlin.de* mit folgenden Angaben auflösen:

- Name, Telefonnummer und E-Mail-Adresse des Verantwortlichen
- Name des zu sichernden Rechners (mit domain)
- Betriebssystem (bei Unix z. B. mit `uname -a` ermittelt)

- Name (mount-point), maximale Grösse und zurzeit belegter Platz der zu sichernden Filesysteme (bei Unix wäre das z. B. das Ergebnis von `df -k` oder einem ähnlichen Kommando)

Wir werden dann mit dem Verantwortlichen in Verbindung treten und ihm die notwendigen Informationen zur Installation zukommen lassen.

- Einen Hinweis zu relevanter Literatur kann auf unserem WWW-Server unter der folgenden URL gefunden werden:

http://www.rz.hu-berlin.de:/rz/filesv/adsm/en_US

Sicherung von VINES-Servern:

- VINES wird durch ADSM nicht unterstützt. Für die Sicherung von VINES-Servern wurde ein Produkt der Firma CommVault beschafft. Nähere Informationen dazu können von unseren VINES-Verantwortlichen angefordert werden.

Christoph Weickmann
weickmann@rz.hu-berlin.de